



## **Plasma metabolites associated with exposure to perfluoroalkyl substances and risk of type 2 diabetes – A nested case-control study**

Downloaded from: <https://research.chalmers.se>, 2023-05-06 06:21 UTC

Citation for the original published paper (version of record):

Schillemans, T., Shi, L., Donat Vargas, C. et al (2021). Plasma metabolites associated with exposure to perfluoroalkyl substances and risk of type 2 diabetes – A nested case-control study. *Environment International*, 146.  
<http://dx.doi.org/10.1016/j.envint.2020.106180>

N.B. When citing this work, cite the original published paper.



# Plasma metabolites associated with exposure to perfluoroalkyl substances and risk of type 2 diabetes – A nested case-control study

Tessa Schillemans<sup>a,\*</sup>, Lin Shi<sup>b,c</sup>, Carolina Donat-Vargas<sup>a,d</sup>, Kati Hanhineva<sup>b,e,f</sup>, Andreas Tornevi<sup>g</sup>, Ingegerd Johansson<sup>h</sup>, Jani Koponen<sup>i</sup>, Hannu Kiviranta<sup>i</sup>, Olov Rolandsson<sup>j</sup>, Ingvar A. Bergdahl<sup>g</sup>, Rikard Landberg<sup>b,g</sup>, Agneta Åkesson<sup>a</sup>, Carl Brunius<sup>b</sup>

<sup>a</sup> Cardiovascular and Nutritional Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden

<sup>b</sup> Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

<sup>c</sup> School of Food Engineering and Nutritional Science, Shaanxi Normal University, Xi'an, China

<sup>d</sup> Department of Preventive Medicine and Public Health, School of Medicine, Universidad Autónoma de Madrid, CEI UAM+CSIC, Madrid, Spain

<sup>e</sup> Institute of Public Health and Clinical Nutrition, University of Eastern Finland, Kuopio, Finland

<sup>f</sup> Department of Biochemistry, University of Turku, Turku, Finland

<sup>g</sup> Department of Public Health and Clinical Medicine, Section of Sustainable Health, Umeå University, Umeå, Sweden

<sup>h</sup> Departments of Odontology, Umeå University, Umeå, Sweden

<sup>i</sup> Department for Health Security, Environmental Health Unit, Finnish Institute for Health and Welfare, Kuopio, Finland

<sup>j</sup> Department of Public Health and Clinical Medicine, Family Medicine, Umeå University, Umeå, Sweden

## ARTICLE INFO

Handling Editor: Shoji F. Nakayama

### Keywords:

Perfluoroalkyl substances  
Metabolomics  
Type 2 Diabetes  
Nested case-control study

## ABSTRACT

Perfluoroalkyl substances (PFAS) are widespread persistent environmental pollutants. There is evidence that PFAS induce metabolic perturbations in humans, but underlying mechanisms are still unknown. In this exploratory study, we investigated PFAS-related plasma metabolites for their associations with type 2 diabetes (T2D) to gain potential mechanistic insight in these perturbations.

We used untargeted LC-MS metabolomics to find metabolites related to PFAS exposures in a case-control study on T2D (n = 187 matched pairs) nested within the Västerbotten Intervention Programme cohort. Following principal component analysis (PCA), six PFAS measured in plasma appeared in two groups: 1) perfluorononanoic acid, perfluorodecanoic acid and perfluoroundecanoic acid and 2) perfluorohexane sulfonic acid, perfluorooctane sulfonic acid and perfluorooctanoic acid. Using a random forest algorithm, we discovered metabolite features associated with individual PFAS and PFAS exposure groups which were subsequently investigated for associations with risk of T2D.

PFAS levels correlated with 171 metabolite features ( $0.16 \leq |r| \leq 0.37$ , false discovery rate (FDR) adjusted p < 0.05). Out of these, 35 associated with T2D (p < 0.05), with 7 remaining after multiple testing adjustment (FDR < 0.05). PCA of the 35 PFAS- and T2D-related metabolite features revealed two patterns, dominated by glycerophospholipids and diacylglycerols, with opposite T2D associations. The glycerophospholipids correlated positively with PFAS and associated inversely with risk for T2D (Odds Ratio (OR) per 1 standard deviation (1-SD) increase in metabolite PCA pattern score = 0.2; 95% Confidence Interval (CI) = 0.1–0.4). The diacylglycerols also correlated positively with PFAS, but they associated with increased risk for T2D (OR per 1-SD = 1.9; 95% CI = 1.3–2.7). These results suggest that PFAS associate with two groups of lipid species with opposite relations to T2D risk.

## 1. Introduction

Perfluoroalkyl substances (PFAS) comprise a large group of chemicals used in a variety of technological applications like textile or leather

treatment, surfactants and food packaging to make them e.g. non-stick and stain repellent. The half-life of PFAS in humans is 3–5 years (Cala-fat et al., 2007; Olsen et al., 2007) and bioaccumulative properties of PFAS increase by chain length and sulfonation with the strongest bio-

\* Corresponding author at: Institute of Environmental Medicine, Karolinska Institutet, Nobels väg 13, Solna, Box 210, SE-171 77 Stockholm, Sweden.

E-mail address: [tessa.schillemans@ki.se](mailto:tessa.schillemans@ki.se) (T. Schillemans).

<https://doi.org/10.1016/j.envint.2020.106180>

Received 5 June 2020; Received in revised form 30 September 2020; Accepted 1 October 2020

Available online 25 October 2020

0160-4120/© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

accumulative properties for long-chain and sulfonated PFAS (Conder et al., 2008). Exposure occurs mainly via consumption of fish and contaminated drinking water (Haug et al., 2010; Stubleski et al., 2017), although house dust and contact with products containing PFAS could contribute (Haug et al., 2011; Vestergren and Cousins, 2009; Strynar and Lindstrom, 2008). PFAS have been hypothesized to influence metabolic control mechanisms involved in the development of type 2 diabetes (T2D) (Sunderland et al., 2019), a major global health problem, characterised by insulin resistance of target tissues and reduced insulin secretion by the pancreatic  $\beta$ -cells (Olokoba et al., 2012).

The chemical structure of PFAS has homology with that of fatty acids and could interfere with fatty acid oxidative or endocrine pathways (Sheng et al., 2018; Kar et al., 2017; Jensen and Leffers, 2008). Foremost, *in vitro* studies show that PFAS are ligands for peroxisome proliferator-activated receptors (PPARs)  $\alpha$ ,  $\gamma$  (Wolf et al., 2008b; Rosen et al., 2017; Vanden Heuvel et al., 2006) and  $\beta/\delta$  with different potencies depending on carbon chain length and terminal functional group (Li et al., 2019; Wolf et al., 2008a). PPARs regulate glucose and lipid metabolism and are involved in adipogenesis, fat storage and inflammation (Bocher et al., 2002; Lefebvre et al., 2006; Poynter and Daynes, 1998). In addition, PFAS may influence  $\beta$ -cell viability (Maillard et al., 2012, 2008), disrupt thyroid hormones (Lee and Choi, 2017) and sex steroid hormone biosynthesis (Zhao et al., 2010; Boujrad et al., 2000; Shi et al., 2010, 2007; Lau et al., 2007), which implies that there may be sex-specific effects. Although some animal data support more of a potential beneficial effect of PFAS on glucose and lipid homeostasis (Yan et al., 2015; Bijland et al., 2011; Kees et al., 1992), the epidemiological evidence is limited and inconsistent, showing direct (He et al., 2018; Sun et al., 2018), null or even inverse associations with T2D (Donat-Vargas et al., 2019a; MacNeil et al., 2009; Conway et al., 2016; Karnes et al., 2014). Our previous study, based on the present study population, assessing PFAS exposures with T2D risk without including the metabolome, found nominal inverse associations of PFAS with insulin resistance and T2D risk, possibly explained by PFAS activation of PPARs (Donat-Vargas et al., 2019a).

However, the relevant metabolic and molecular pathways involved in T2D development that may be affected by PFAS have not been fully elucidated. Thus, to gain more insight in the PFAS-T2D relation, metabolomics can be a useful tool (Wild, 2005; Rattray et al., 2018) as the presence and concentration of metabolites directly reflects the underlying biochemical activity (Kim et al., 2016). So far, few studies have investigated metabolites associated with PFAS (Salihovic et al., 2019), of which two were in children (Alderete et al., 2019; Kingsley et al., 2019) and one related to pregnancy (Hu et al., 2019). These studies indicated mainly perturbations in lipid and fatty acid metabolism and also found distinctive metabolic profiles for different PFAS (Salihovic et al., 2019), nevertheless, none investigated associations with T2D risk.

Thus, we analysed untargeted metabolomics data from a T2D case-control study nested in a population-based prospective cohort. Our objectives were to discover metabolites related to PFAS exposures, explore potential differences in metabolic profiles of six chemically different PFAS, and to investigate whether PFAS-related metabolite patterns were associated with risk of developing T2D.

## 2. Materials and methods

### 2.1. Study population

This nested case-control study from the Västerbotten Intervention Programme cohort included participants that had donated overnight fasting blood samples to the biobank on two occasions on average 10 ( $\pm 4.4$ ) years apart (of which at least the first was prior to the case diagnosis) (Donat-Vargas et al., 2019a). The Västerbotten Intervention Programme, initiated 1985, invited all residents in Västerbotten County in Northern Sweden for health screening and blood sampling when they turned 40, 50 and 60 years. This study includes baseline data 1990–2003

and follow-up data 2000–2013. Participation rate exceeded 56%, being often around 70%, and 90.5% also donated blood samples (Norberg et al., 2010). Incident diabetes cases were identified in the DiabNorth register (last update in 2008), which had a participation rate of 70% of VIP participants diagnosed with diabetes, (Rolandsson et al., 2012) and were diagnosed according to WHO recommendations and analysis of autoantibodies (Rolandsson et al., 2012). The average time from baseline blood sampling to diabetes diagnosis was 8.2 years. Diabetes cases were individually matched with controls based on age, sex and date of blood draw. PFAS and metabolites were measured in 374 participants at baseline and follow-up (187 diabetes case-control pairs). All 374 participants were included in an initial exploratory phase where PFAS-related metabolite features were discovered, as indicated in Fig. 1. For the association analyses between PFAS-related metabolite features and T2D, analysis was restricted to 124 T2D case-control pairs (excluded matched pairs if cases had an unclear diabetes diagnosis ( $n = 34$  pairs) or had diabetes at baseline ( $n = 21$  pairs) and if controls had possible diabetes ( $n = 8$  pairs)) (Donat-Vargas et al., 2019a).

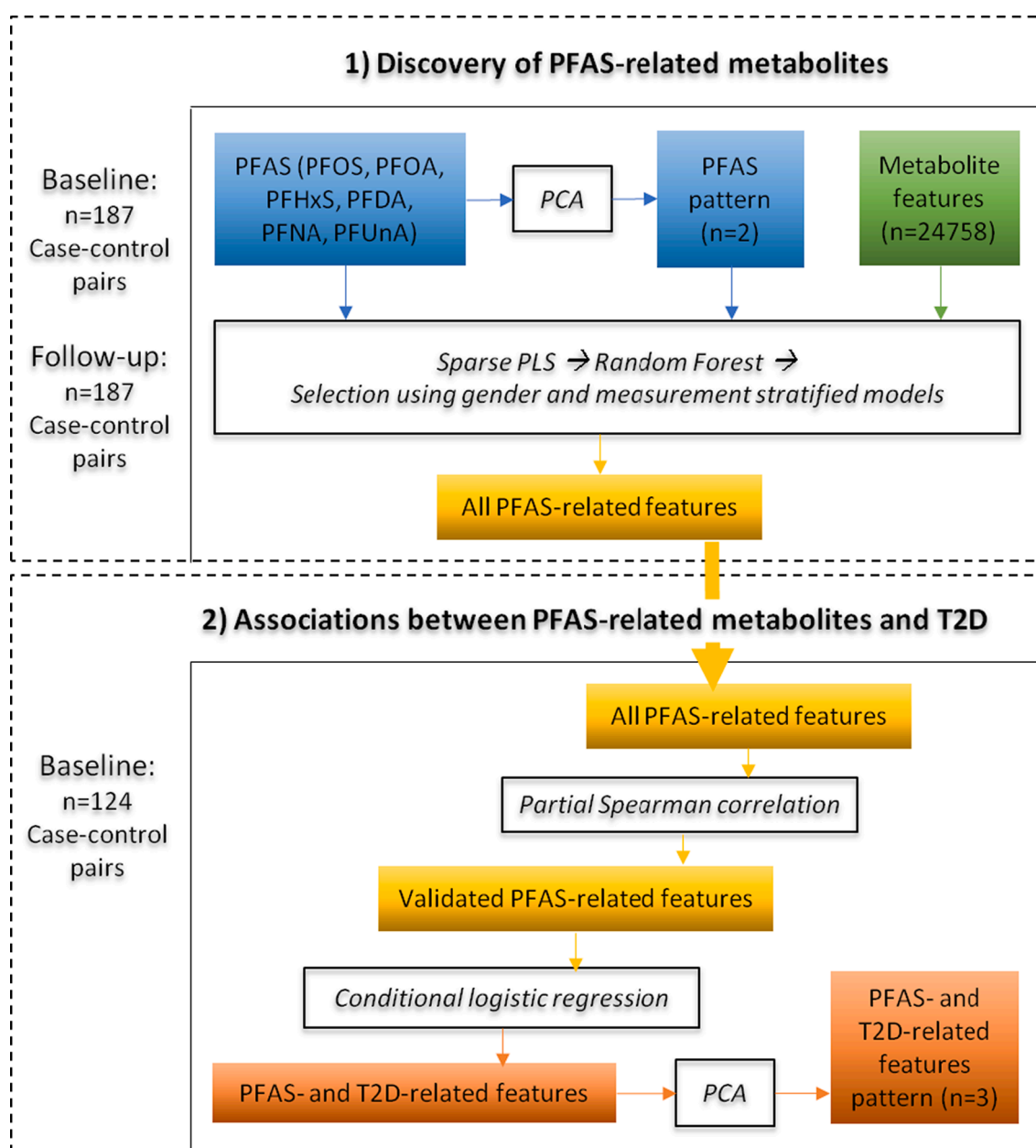
Participation information, collected via questionnaires, included marital status, attained education, smoking status, body mass index (BMI) and physical activity (as determined by Cambridge index for physical activity). Based on a validated food frequency questionnaire, we retrieved information on red/processed meat, fish (g/day per 1000 kcal) and alcohol consumption (g/day) (Johansson et al., 2001, 2002; Nettleton et al., 2013). Informed consent was obtained from all participants. The study protocol was approved by the regional ethics review boards in Umeå and Uppsala, Sweden (Dnr 2013/414-31, 2014/147-32 M and 2014/011).

### 2.2. Targeted LC-MS PFAS measurements

PFAS concentrations were measured in plasma samples from baseline and follow-up through targeted liquid chromatography-triple quadrupole mass spectrometry (LC-MS/MS) as previously described (Koponen et al., 2013). Six different PFAS compounds had detectable concentrations: perfluorohexane sulfonic acid (PFHxS), perfluorooctane sulfonate (PFOS), perfluorooctanoic acid (PFOA), perfluorononanoic acid (PFNA), perfluorodecanoic acid (PFDA) and perfluoroundecanoic acid (PFUnDA). Concentrations below the limit of quantitation (LOQ) were set to the  $LOQ/2 = 0.075$  ng/mL for 26% of the subjects in PFDA and 42% in PFUnDA (Donat-Vargas et al., 2019a; Richardson and Ciampi, 2003).

### 2.3. Untargeted LC-MS metabolite profiling

Untargeted LC-quadrupole time of flight (qTOF)-MS metabolic profiling was performed on baseline and follow-up plasma samples (Shi et al., 2018b). Data acquisition, corrections and normalizations were performed previously and have been described elsewhere (Brunius et al., 2016; Shi et al., 2017). In brief, de-proteinised fasting heparin plasma samples were analysed by LC-qTOF-MS (Agilent Technologies, United States) on reverse phase and hydrophilic interaction liquid chromatography columns in both positive and negative ionisation modes. The MassHunter Acquisition B.04.00 software (Agilent Technologies) was used for data acquisition. Raw data was converted to mzXML format and deconvolution was performed ('XCMS' R package) (Shi et al., 2017). The stability of the instrumental analysis was monitored by quality control samples (batch-specific pooled aliquots of the study samples and a long-term reference sample were injected at the beginning, end and regularly throughout the injection sequence). Measurement drift was calculated per batch by a cluster-based approach using batch-specific quality control samples with signal correction performed per cluster if resulting in improved measurement homogeneity among long-term reference samples ('batchCorr' R package) (Brunius et al., 2016). Features were over-all retained if they passed QC test ( $CV < 0.3$ ) in at least five out of the eight analytical batches (Supplementary Material of L. Shi et al.



**Fig. 1.** Flow chart of the analytical study design divided in selection of PFAS-related features (top) and associations between PFAS-related features and risk of T2D (bottom). PFAS-related features were discovered using the random forest algorithm in step 1 and validated in step 2 for the reduced sample population at baseline after inclusion of confounders. Subsequently, the validated features were investigated for their prospective associations with T2D. Abbreviations: Principal Component Analysis (PCA) and Partial Least Squares (PLS).

2018 (Shi et al., 2018a)). Missing values were replaced with values selected at random from a normal distribution between 0 and the lowest measured peak intensity for each metabolite feature under the assumption of missing not at random due to low concentration. Poorly retained lipids in hydrophilic interaction chromatography (retention time < 70 s) were removed, as these were better represented in reverse phase. We use the term metabolite feature to refer to a molecular entity with a unique mass to charge ratio and retention time as measured by LC-MS. We putatively annotated six metabolite features as possible PFAS (Supplementary Table 1), which were excluded from the metabolomics dataset in order not to bias data models by being included both as independent and dependent variables.

#### 2.4. Statistical analysis

A flow-chart of the analytical procedure is presented in Fig. 1. To discover metabolite features related not only to individual PFAS exposures, but also PFAS exposure groups, we performed an orthogonally (varimax) rotated principal component analysis (PCA) on individual square root-transformed PFAS scaled to unit variance. We obtained two PFAS groups through manual inspection of the first two principal component (PC) loadings (eigenvalue > 1). The two PFAS groups and the individual PFAS were later tested for their correlations with PFAS-related metabolite features, to investigate whether the PFAS groups associated with similar metabolic alterations as the individual PFAS. We filtered out the majority of metabolite features likely not associated with PFAS by sparse partial least squares regression using four components and the metabolomics data as the predictor variables and each of

individual square root-transformed PFAS concentrations and two PFAS groups as response variables. The arbitrary choice of four components allowed for a compromise between modelling complexity and parsimony and reduced the number of features from 24,758 to approximately 4000 and was performed primarily to improve the computational performance in the subsequent predictive modelling: The pre-filtered metabolomics data were entered as predictor data and processed using a random forest model within a repeated double-cross validation framework incorporated with unbiased variable selection (R package MUVR (Shi et al., 2019a), tutorial available on <https://gitlab.com/CarlBrunius/MUVR>), using each of the individual square root-transformed PFAS concentrations and PFAS groups ( $n = 2$ ) as response variables. In order to assess modelling performance, permutation analysis was performed ( $n = 50$ ,  $p < 0.001$ ) (Lindgren et al., 1996), parameter settings for the random forest models and permutation analyses are described in Supplementary Table 2. Moreover, to exclude metabolite features that likely originated from sex or baseline versus follow-up measurement differences, we tested the robustness of the discovered metabolite features associated with PFAS by a series of similar analyses, stratified by sex and measurement occasion. Those metabolite features that were present both in the overall model and in at least one of the sex and measurement stratified models, were further validated using partial Spearman correlation with PFAS levels at the baseline measurement while adjusting for the following potential confounding factors: sex, age, sample year, marital status, education, smoking status, physical activity and case-control status. PFAS were square root-transformed and metabolite features were square root-transformed as well as standardized prior to the correlation analysis and p-values were False Discovery Rate adjusted for multiple testing ( $FDR < 0.05$ ).

In the second step, PFAS-related metabolite features were tested for their individual prospective associations with T2D risk using conditional logistic regressions, with standardized square root-transformed metabolite features as the independent variables. In model 1, the analyses were adjusted for age, sex and sample year, marital status, education, smoking status and physical activity, and in model 2 additionally for fish-, red/processed meat- and alcohol consumption and BMI (Donat-Vargas et al., 2019a). Resulting p-values from all models were adjusted for multiple testing ( $FDR < 0.05$ ). Odds ratio (OR) is presented per 1-standard deviation (SD) increase. The likelihood for discovering T2D-related metabolite features amongst the PFAS-related metabolite features was assessed using the hypergeometric test ( $p < 0.05$ ).

In order to investigate intercorrelations between PFAS- and T2D-related metabolite features, a PCA was constructed based on the 35 selected metabolite features associated with T2D ( $p < 0.05$ ). Correlations between PCs ( $n = 3$ , eigenvalue  $> 2$ ) and individual PFAS, PFAS groups, food intakes (i.e. fish, red/processed meat), alcohol consumption, BMI and HOMA-IR were analysed using partial Spearman correlation analysis adjusted for sex, age, sample year, marital status, education, smoking status, physical activity and case-control status. Associations between PC scores and T2D risk were analysed using conditional logistic regression adjusted for covariates according to model 1 and model 2 (as above). In a sensitivity analysis, we also investigated associations between PC scores and T2D risk while stratifying for sex. OR is presented per 1-SD increase in the pattern score. The associations between T2D- and PFAS-related metabolite feature patterns, the PFAS levels and diet with T2D risk were displayed in a triplot (Schillemans et al., 2019) (R package Triplot, tutorial available on <https://gitlab.com/CarlBrunius/triplot>). Briefly, the triplot is a visualization tool that facilitates interpretation of exposures, the metabolome and disease risk data. It presents loadings of metabolite features in 2 PCs (in black), whilst superimposing the correlations between the PCs and the PFAS levels or diet components (in blue) as well as the associations between the PCs and T2D risk (in red). Thus, metabolites, PFAS levels/diet components and T2D risk pulling in the same axis in the triplot figure are positively correlated/directly associated, whereas opposite axis are

negatively correlated/inversely associated.

These findings were validated in a second dataset containing participants from the same case-control study with similar background characteristics (Supplementary Table 3), but with only one measurement (baseline) and without PFAS measurements ( $N = 598$ ) (Shi et al., 2018a). PC scores for the PFAS and T2D-related metabolite feature patterns were predicted in the validation dataset based on the PCA of the original dataset (described above) and then used in a conditional logistic regression with T2D risk.

To avoid confusion, the term “group” refers to exposure groups obtained from the PCA on PFAS, whereas the term “pattern” refers to metabolite feature patterns obtained from the PCA on the PFAS- and T2D-related metabolite features. All statistical analyses were performed using R software version 3.4.3 (R Core Team, Vienna, Austria) (R Core Team, 2017).

## 2.5. Metabolite identification

The metabolite annotation is reported in Supplementary Table 4, following the Metabolomics Standards Initiative (MSI) reporting criteria for the confidence level. Auto-MS/MS data was only available for a few of the selected metabolite feature peaks, some of these could be matched to literature based on accurate mass and product ion spectrum (level 2) (Shi et al., 2018b, 2019b; Hanhineva et al., 2014). Other metabolite features associated with PFAS and T2D were putatively annotated for compound class based on  $m/z$  (mass tolerance  $< 20$  ppm) and retention time using matching against online databases (level 3). Unknown compounds were presented as “analytical mode  $_m/z$  @ retention time” (level 4).

## 3. Results

### 3.1. Population and PFAS

Total study population characteristics at baseline are described in Table 1.

Differences in characteristics and PFAS concentrations between cases and controls at baseline and follow-up as well as the PFAS intercorrelations are described in our previous papers (Donat-Vargas et al., 2019a, 2019b). In brief, T2D cases had somewhat higher BMI and physical activity, but lower PFAS concentrations (except for PFNA) compared to the controls (Donat-Vargas et al., 2019a). High correlations were found between PFOS and PFOA as well as between PFNA, PFDA

**Table 1**  
Baseline main characteristics of the total study population ( $N = 374$ ).

Characteristics	Total Study Population ( $N = 374$ )
Females <sup>a</sup> , n (%)	172 (46)
Age <sup>a</sup> (years)	46 $\pm$ 6
Body Mass Index (kg/m <sup>2</sup> )	27 $\pm$ 4
Education, n (%)	
>12 yrs.	270 (72)
Smoking status, n (%)	
Former	144 (39)
Current	86 (23)
Physical activity <sup>b</sup> , n (%)	
Inactive	201 (54)
Fish (g/day/1000 kcal)	10.0 $\pm$ 7.8
Meat (g/day/1000 kcal)	50.3 $\pm$ 21.2
Alcohol consumption, n (%)	
0.1–5 (g/day)	260 (70)
5–15 (g/day)	87 (23)
>15 (g/day)	6 (2)

Note: Continuous variables are given as mean  $\pm$  SD and categorical variables are given as number (percentage).

<sup>a</sup> Matching factors for case-control pairs.

<sup>b</sup> Cambridge index for physical activity.



and PFUnDA (Donat-Vargas et al., 2019a). In addition to the previously observed results where PFOS and PFOA had higher concentrations at baseline, whereas PFNA and PFDA seemed higher at follow-up (Donat-Vargas et al., 2019b), a tendency for higher PFAS concentrations was observed in males compared to females, with exception of PFUnDA (Fig. 2). The PCA of the six PFAS showed that two PCs accounted for 73% of the variance. High loadings were observed for the long-chain PFDA, PFNA and PFUnDA in PC1 (henceforth referred to as LC\_PFAS) and for the shorter-chain PFHxS, PFOS and PFOA in PC2 (SC\_PFAS) (Supplementary Table 5).

### 3.2. PFAS-related metabolite features

The number of metabolite features observed per individual PFAS and PFAS group after each analysis step is presented in Table 2. We found 290 unique metabolite features related to one or more PFAS in the multivariate random forest models (permutation analysis  $p < 0.001$ ), after filtering out those features likely to be selected because of sex or time of measurement differences alone (Supplementary Table 6). Out of these 290 metabolite features, 171 correlated with individual PFAS at baseline in the selected population ( $n = 124$ ) after adjustment for confounders ( $0.16 \leq |r| \leq 0.37$ ,  $FDR < 0.05$ ). Among them, a majority of metabolite features were associated with PFNA, PFDA, PFUnDA or their affiliated LC\_PFAS group (Table 2).

### 3.3. PFAS and T2D-related metabolite features

Among the 171 PFAS-associated metabolite features, 35 associated with T2D adjusted for matching factors, marital status, education, smoking status and physical activity in model 1, (Supplementary Fig. 1), but only 13 remained after full multivariable-adjustment in model 2 (model 1 additionally adjusted for fish-, red/processed meat- and alcohol consumption and BMI) (Fig. 3). After adjustment for multiple testing, 7 metabolite features remained in model 1 (hypergeometric  $p < 0.0025$ ; Supplementary Table 7–8), but only 1 in model 2 (hypergeometric  $p = 0.45$ ; Supplementary Table 7–8). One metabolite annotated

**Table 2**

Number of metabolite features (maximum model) related per individual PFAS or PFAS pattern exposures and model performance by prediction fitness and permutation analysis ( $N = 374$ ).

Number of metabolite features		Model Performance Prediction fitness <sup>a, b</sup>		Number of metabolite features after filtering <sup>c</sup>	Number of metabolite features after correlation <sup>d</sup>
		R <sup>2</sup>	Q <sup>2</sup>		
<i>Individual PFAS</i>					
PFOS	150	0.415	0.347	36	19
PFOA	150	0.339	0.234	45	14
PFHxS	140	0.417	0.336	35	18
PFDA	220	0.376	0.307	81	66
PFNA	200	0.385	0.315	72	62
PFUnDA	180	0.334	0.243	78	59
<i>PFAS pattern<sup>e</sup></i>					
LC, PFAS	250	0.42	0.357	82	65
SC, PFAS	100	0.465	0.38	27	13

<sup>a</sup> R<sup>2</sup> indicates goodness of fit of the model. Q<sup>2</sup> indicates predictive ability of the model.

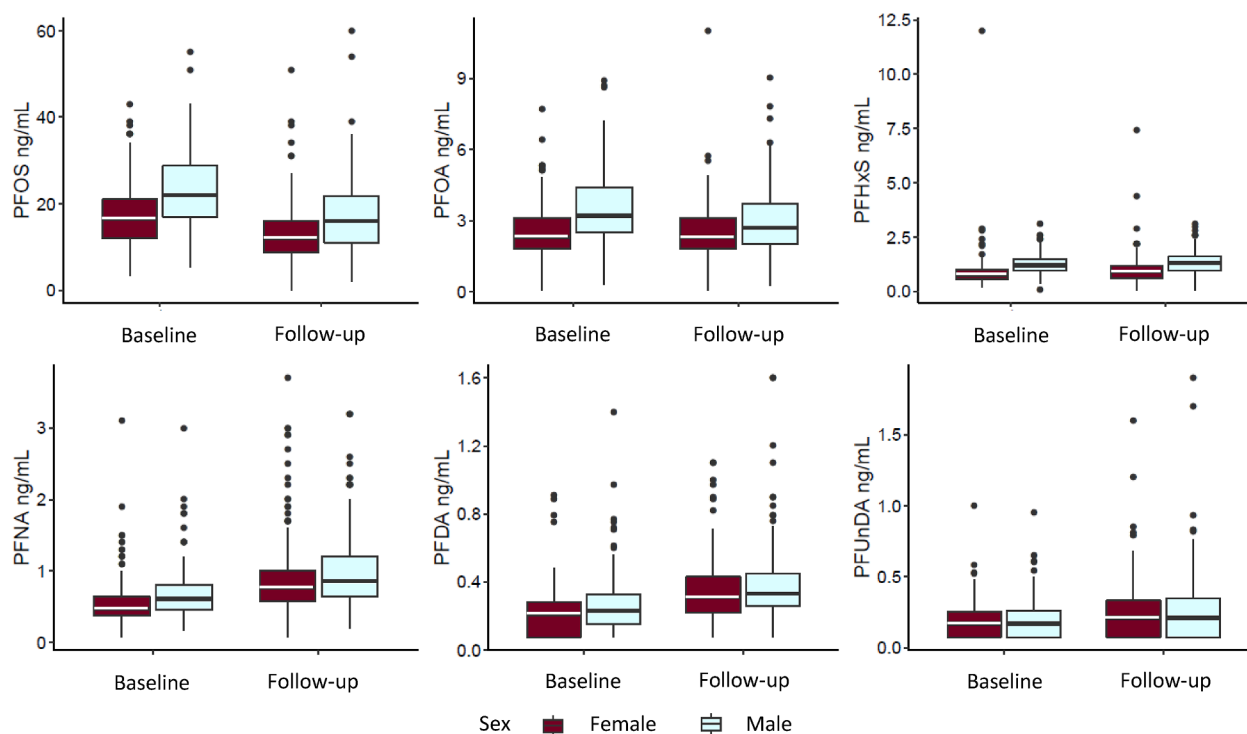
<sup>b</sup> Permutation  $p < 0.001$  from permutation analysis ( $n = 50$ ) for weakest two models (PFOA and PFUnDA).

<sup>c</sup> I.e. metabolite features selected from the full sample model (males, females and both measurements) that also occur in at least one of the sex and measurement time stratified models (i.e. males and females baseline only; males and females follow-up only; males both measurements; males baseline only; males follow-up only; females both measurements; females baseline only; females follow-up only).

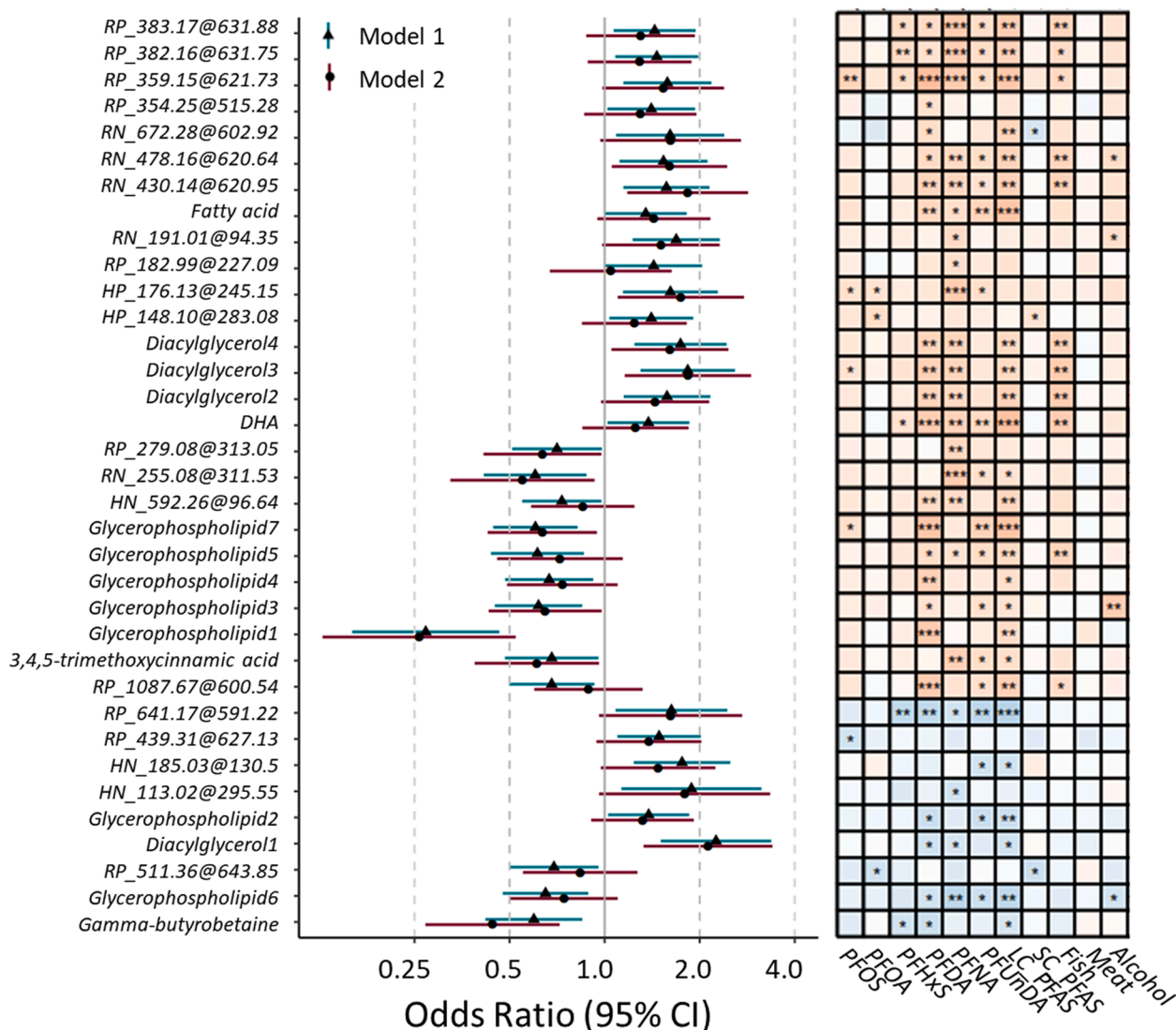
<sup>d</sup> I.e.  $FDR < 0.05$  after partial correlation analysis adjusted for sex, age, sample year, marital status, education, smoking status, physical activity and case-control status ( $n = 248$ ).

<sup>e</sup> Refers to long-chain PFAS: PFDA, PFNA and PFUnDA (LC\_PFAS) and short-chain PFAS: PFHxS, PFOS and PFOA (SC\_PFAS).

as gamma-butyrobetaine (GBB) correlated negatively with PFAS but associated inversely with T2D risk. Other features with strong associations to both PFAS and T2D risk were annotated to compound classes of



**Fig. 2.** PFAS concentrations (ng/mL) in the entire study population ( $n = 374$ ) at baseline and follow-up for males and females. All PFAS, except PFUnDA, had higher concentrations in males compared to females.



**Fig. 3.** Multivariable-adjusted univariate associations between metabolite features and (a) T2D risk (odds ratios per 1-standard deviation increase  $\pm$  confidence interval (CI)) and (b) their correlations with PFAS and food components (missing  $n = 2$ ). (a) Model 1 was adjusted for matching factors (sex, age and sample year), marital status, education, smoking status and physical activity. Model 2 was additionally adjusted for fish, meat and alcohol intake and BMI (missing  $n = 2$ ). (b) Partial Spearman correlations were adjusted for sex, age, sample year, marital status, education, smoking status, physical activity and case-control status and significance is indicated by \* FDR < 0.05, \*\* FDR < 0.01, \*\*\* FDR < 0.001. Red color represents positive correlations, blue color represents negative correlations. Unknown metabolite features are presented as: Mode  $m/z$ @RT. Abbreviations: docosahexaenoic acid (DHA), long-chain PFAS: PFDA, PFNA and PFUnDA (LC\_PFA) and short-chain PFAS: PFHxS, PFOS and PFOA (SC\_PFA). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

glycerophospholipids and diacylglycerols.

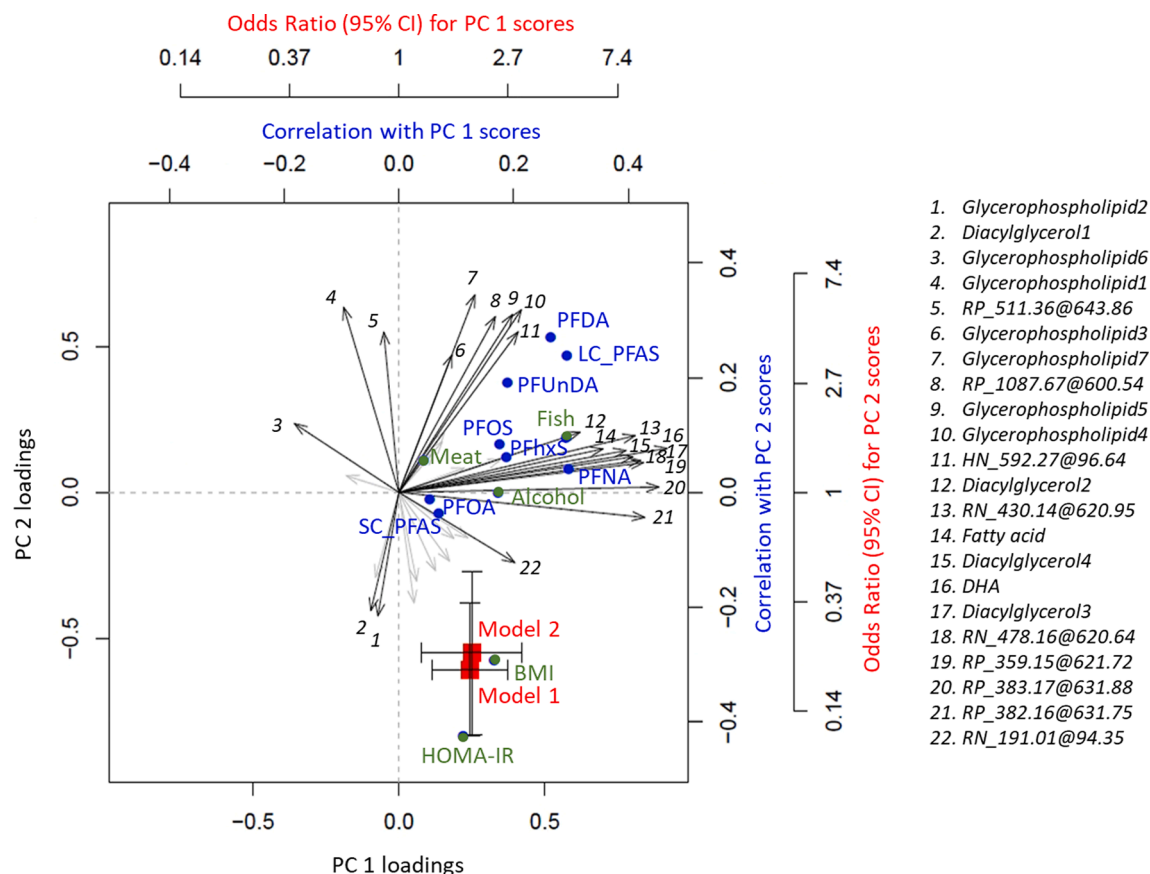
In the PCA on the 35 metabolite features selected for their associations with PFAS and diabetes (Supplementary Fig. 2, Supplementary Table 5), PC1 and PC2 accounted for the largest proportion of the total variance (23% and 12%, respectively) and showed the strongest association with T2D risk, thus these were further investigated. We found that metabolite features putatively annotated as glycerophospholipids had high loadings in PC2, whereas those annotated as diacylglycerols had high loadings in PC1 (Supplementary Table 5, Fig. 4). PFDA, PFUnDA and LC\_PFA correlated positively with both PC1 and PC2, whereas PFNA correlated only with PC1. PFHxS, PFOS, PFOA and SC\_PFA did not correlate with any of the PCs. PC1 furthermore correlated positively with alcohol and fish intake as well as with BMI and HOMA-IR and was associated with increased risk of T2D (model 1: OR = 1.9; 95% Confidence Interval (CI) = 1.3–2.7). PC2 correlated negatively with BMI and HOMA-IR, but not with any of the dietary intakes and associated with lower risk of T2D (model 1: OR = 0.2; 95% CI = 0.1–0.4)

(Fig. 4).

Mutual adjustment for PC1 and PC2 did not affect results qualitatively (data not shown). Similar associations with T2D were seen in the validation cohort (Supplementary Fig. 3). Results did not differ greatly between males and females, although males had a slightly higher OR of T2D for PC1, whereas females had a higher OR for PC2 (Supplementary Fig. 3).

#### 4. Discussion

This population-based prospective nested case-control study is the first to use metabolomics data to examine the relationship between PFAS and risk of T2D, providing valuable information about potential underlying mechanisms. Modelling of PFAS exposures by the metabolome in the random forest algorithm resulted in strikingly strong models and 171 metabolite features associated with PFAS, from which 35 also associated with T2D risk. This is more than would be expected by chance



**Fig. 4.** Multivariate associations of PFAS and T2D-related metabolite features with PFAS, fish and red/processed meat intakes, alcohol consumption, BMI, HOMA-IR and risk of T2D. The triplot represents the correlations between metabolite features (features with absolute loading > 0.4 are visualized by black arrows with label, whereas with absolute loading < 0.4 are represented by light grey arrows), fish, red/processed meat intakes, alcohol consumption, BMI and HOMA-IR as well as risk of T2D (Schillemans et al., 2019). Correlations were calculated using partial Spearman correlation adjusted for sex, age, sample year, case/control status, marital status, education, smoking status, physical activity and FDR multiple testing (missing  $n = 2$ ). Odds ratios per 1 standard deviation increase in pattern score  $\pm 95\%$  confidence intervals were obtained from two conditional logistic regression models: Model 1 was adjusted for matching factors (sex, age and sample year), marital status, education, smoking status and physical activity. Model 2 was additionally adjusted for fish, meat and alcohol intake and BMI (missing  $n = 2$ ). Unknown metabolite features are presented as: Mode  $m/z$ @RT. Abbreviations: docosahexaenoic acid (DHA), long-chain PFAS: PFDA, PFNA and PFUnDA (LC\_PFAS) and short-chain PFAS: PFHxS, PFOS and PFOA (SC\_PFAS). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

( $n = 2$ ), indicating that most of the 35 features could be involved in pathways related to both PFAS and T2D. Such common pathways are of interest for the understanding of biological mechanisms behind any relationship between PFAS exposures and T2D risk.

PFAS were, via PCA, assembled into two groups based on their measured concentrations in blood, with shorter-chain PFHxS, PFOS and PFOA in one group (i.e. SC\_PFAS) and long-chain PFNA, PFDA and PFUnDA in the other (i.e. LC\_PFAS). However, the underlying rationale of the observed grouping is uncertain. PFAS could potentially group together either due to similarities in exposure sources or to their biological accumulation and potency (Wolf et al., 2008a). Several lipid-related metabolites were associated with individual PFAS congeners or PFAS groups. These findings are supported by previous untargeted metabolomics studies, which also indicate associations between PFAS (i.e. PFHxS, PFOS, PFOA, PFNA, PFDA, PFUnDA) and fatty acids, lipids and, in particular, glycerophospholipids (Salihovic et al., 2019; Alderete et al., 2019; Kingsley et al., 2019; Hu et al., 2019). The current results are further strengthened by the fact that similar PFAS associated metabolites were discovered using different data analytical strategies, i.e. univariate analysis (Salihovic et al., 2019; Alderete et al., 2019; Kingsley et al., 2019; Hu et al., 2019) versus the multivariate modelling introduced in this work. Although causality cannot be inferred in the present study design, a PFAS-related effect on lipids is biologically plausible

since PFAS are thought to disrupt fat metabolism through several mechanisms, e.g. activation of PPARs (Fievet et al., 2006; Berger et al., 2005), interference with endocrine pathways (Zhao et al., 2010) or cholesterol transport (Boujrad et al., 2000; Shi et al., 2010, 2007).

Unsupervised analysis of the PFAS-associated metabolite features showed that one pattern was dominated by annotated glycerophospholipids and the other by diacylglycerols. Both PFAS-related metabolite feature patterns correlated with the long-chain PFNA, PFDA and PFUnDA, whereas they did not correlate equally strong with the shorter PFHxS, PFOS and PFOA. PFDA and PFUnDA were mainly associated with the glycerophospholipid metabolites, whereas PFNA associated with diacylglycerol metabolites. The stronger associations for long-chain PFAS may be explained by a higher potency and biological accumulation (Wolf et al., 2008a), supported by a previous finding where the strongest associations were also found between long-chain PFAS (i.e. PFNA and PFUnDA) and glycerophospholipids as well as other fatty acids (e.g. docosapentaenoic acid (DPA) and docosahexaenoic acid (DHA)) (Salihovic et al., 2019).

Furthermore, we discovered that the two metabolite feature patterns related to long-chain PFAS had opposite associations with T2D risk. The glycerophospholipid pattern associated with a decreased risk of T2D, whereas the diacylglycerol pattern associated with an increased risk. This is in line with a previous plasma lipidomic profiling study that



showed inverse associations between certain phospholipids, especially lyso-phosphatidylcholines, and T2D as well as direct associations between diacylglycerols and T2D (Razquin et al., 2018). Our results suggest that PFAS are associated with two groups of lipid species with opposite relations with T2D development, which could contribute to the interpretation of inconsistent results from epidemiological studies investigating PFAS exposures and subsequent T2D risk. A prospective and a cross-sectional study indicate associations of PFOS and PFOA with increased T2D risk (He et al., 2018; Sun et al., 2018), which would be supported by the results of the diacylglycerol dominated metabolite feature pattern. On the other hand, a few cross-sectional studies show lower T2D prevalence with higher serum PFAS (i.e. PFHxS, PFOS, PFOA and PFNA) concentrations (MacNeil et al., 2009; Conway et al., 2016). Likewise, in this same study population, we previously showed inverse prospective associations between PFAS (i.e. PFHxS, PFOS, PFOA, PFNA, PFDA and PFUnDA) levels and T2D risk as well as inverse associations with insulin resistance and triglycerides in those who were non-diabetics, although these were mainly non-significant (Donat-Vargas et al., 2019a, 2019b). These findings get support from the results of the glycerophospholipid dominated metabolite feature pattern, which also shows an inverse correlation with HOMA-IR. Thus, PFAS-related glycerophospholipids may be related to insulin sensitivity and T2D risk. Interestingly, glycerophospholipids and diacylglycerols share reactions in lipid synthesis pathways (Saltiel and Kahn, 2001). Additionally, increased circulating free fatty acids and accumulation of triglycerides and fatty acid-derived metabolites (e.g. diacylglycerols) in muscle and liver is thought to contribute to insulin resistance (Saltiel and Kahn, 2001).

Another potentially interesting PFAS- and T2D-related metabolite was identified as GBB, an intermediate metabolite in the carnitine biosynthesis and in the gut microbiota-dependent formation of trimethylamine N-oxide (Koeth et al., 2019). GBB correlated negatively with PFAS and associated inversely with T2D. One possible pathway linking PFAS to GBB relates to PPAR $\alpha$  activation, as animal studies indicate that carnitine metabolism may be regulated by PPAR $\alpha$  (Ringseis et al., 2012), another speculation is that PFAS affects the gut microbiota. Carnitine plays a key role in lipid metabolism and studies indicate both beneficial (Strand et al., 2018) and detrimental (Koeth et al., 2014) effects of carnitine on T2D due to fatty acid transport and  $\beta$ -oxidation as well as trimethylamine N-oxide associated pro-diabetic effects (e.g. inflammation, insulin resistance and hyperlipidemia) (Li et al., 2015; Dambrova and Liepinsh, 2015). Red meat is the main source for carnitine, but there was no correlation between GBB and meat or fish, making it unlikely that the PFAS-GBB link is confounded by diet.

Nevertheless, since different dietary habits and different dietary sources of PFAS have been suggested to confound the PFAS-T2D association and contribute to the inconsistencies in the epidemiological findings (Donat-Vargas et al., 2019a), we further explored correlations between other PFAS-related metabolite features and different dietary components. In this study, fish intake correlated the strongest with the diacylglycerol dominated metabolite feature pattern associated with increased T2D risk. For fish intake, literature indicates inverse associations with T2D (Jannasch et al., 2017), although there are some inconsistencies (Namazi et al., 2019). Since fish is known to be the main dietary source of PFAS as well as of other persistent organic pollutants (POPs) and the latter are associated with increased T2D risk (Tornevi et al., 2019), we speculate that the positive association between PFAS and T2D may be confounded by other POPs. This is supported by our previous finding that some of the PFAS-related metabolites including diacylglycerol (RP\_684.55@805.11) and a known biomarker of fish intake DHA (RP\_329.25@632.04) were also related to fish intake and POPs (Shi et al., 2019b). Although BMI correlated with the PFAS-related metabolite feature pattern, adjustment for BMI in the logistic regression did not affect the qualitative assessment of risk. This finding suggests that the associations between the PFAS-related metabolite features and T2D risk were independent of BMI. It is uncertain whether we should

consider BMI as a confounder or a mediator, as PFAS exposures may affect weight and weight regain (Liu et al., 2018; Cardenas et al., 2018).

Our study fits in a broader perspective of exposome-based studies that are using metabolomics in population-based research as a tool to investigate the influence of environmental exposures on disease risks (Walker et al., 2019). We employed a meet-in-the-middle approach, where the exposure is related to intermediate biomarkers (metabolomics), which are then connected to disease risk (Vineis and Perera et al., 2007; Chadeau-Hyam et al., 2011; Gängler et al., 2019). This approach has advantages for causality assessment in epidemiology and may be integrated with the toxicological concept of adverse outcome pathways. Our study has several strengths. First, the prospective design, measurements of PFAS in blood and stringent T2D diagnosis reduced the likelihood of reverse causation bias and chances of exposure and outcome misclassification, respectively. Second, the collected data allowed adjustments for several potential confounders as well as to explore the impact of diet on the associations. Finally, the random forest modelling was designed to minimize overfitting and false positives while selecting all-relevant metabolite features, allowing us to select metabolic features using a multivariate approach. Nevertheless, there are also several limitations that need to be acknowledged. First, the explorative nature of our study hinders inference regarding possible biomarkers or dose-response relations. Our study may also be subject to false-negative results, i.e. we may have missed associations between PFAS and metabolite features due to stringent inclusion criteria and stratification analyses, as well as false-positive results due to chance findings or overfitting. However, the hypergeometric test indicated a significant enrichment in our results in model 1, but not in model 2, which may be due to over-adjustment in model 2. The validation in the second independent dataset indicates that overfitting is limited, although the possibility for residual confounding remains. Second, we focused on 2 PCs of 35 PFAS-related metabolite features that explained approximately a third of the variance within the data. It is possible that remaining variability might be attributed to either residual variability or potentially important but unknown systematic variability. However, the primary aim of this study was to identify plasma metabolites associated with exposure to both PFAS and T2D risk, which were reflected by the first 2 PCs and were thus further investigated. There was no opportunity to replicate these results in a separate cohort, but similarities in results between different metabolomics studies investigating PFAS do strengthen the observed findings. Furthermore, not all the discovered PFAS-related metabolite features had MS-MS data available and were thus only putatively annotated for class, i.e. glycerophospholipids and diacylglycerols. There was no possibility to do targeted MS-MS yet, but detailed IDs will be included upon targeted examinations.

In conclusion, we annotated metabolite classes that were associated both with baseline PFAS levels and prospective T2D risk in this exploratory nested case-control study. These metabolite features formed two patterns mainly represented by annotated glycerophospholipid and diacylglycerol classes. Although both PFAS-related metabolite feature patterns correlated positively with PFAS, they had opposite directions on T2D risk, with glycerophospholipids showing inverse associations and diacylglycerols showing direct associations. Furthermore, GBB correlated negatively with PFAS and associated inversely with T2D. Our findings highlight the possibility that PFAS are associated with two groups of lipid species with diverging relations with T2D development, which could explain the conflicting associations found in literature between PFAS and T2D risk (Sun et al., 2018; Donat-Vargas et al., 2019a).

#### CRedit authorship contribution statement

**T. Schillemans:** Methodology, Formal analysis, Visualization, Writing - original draft. **L. Shi:** Data curation, Software. **C. Donat-Vargas:** Data curation, Supervision. **K. Hanhineva:** Resources, Data curation. **A. Tornevi:** Data curation. **I. Johansson:** Data curation. **J. Koponen:** Resources. **H. Kiviranta:** Resources. **O. Rolandsson:**

Resources. **I.A. Bergdahl:** Funding acquisition, Methodology. **R. Landberg:** Funding acquisition, Methodology. **A. Åkesson:** Conceptualization, Funding acquisition, Methodology, Supervision. **C. Brunius:** Conceptualization, Funding acquisition, Methodology, Software, Supervision.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgement

We thank the Biobank Research Unit at Umeå University, Västerbotten Intervention Programme, and the County Council of Västerbotten for providing data and samples and acknowledge the contribution from Biobank Sweden, supported by the Swedish Research Council (VR 2017-00650) and the Northern Sweden Diet Database and the funds supporting it. We further acknowledge the EU Horizon 2020 project LongITools (grant agreement 874739). This work has received funding from the Swedish Research Council (VR 2017-05840 and VR 2017-00822), The Swedish Research Council for Sustainable Development (FORMAS 2016-00314), the Swedish Research Council for Health, Working Life and Welfare (FORTE 2012-0758) and the Dr P. Håkansson Foundation.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envint.2020.106180>.

## References

- Alderete, T.L., et al., 2019. Perfluoroalkyl substances, metabolomic profiling, and alterations in glucose homeostasis among overweight and obese Hispanic children: A proof-of-concept analysis. *Environ. Int.* 126, 445–453.
- Berger, J.P., Akiyama, T.E., Meinke, P.T., 2005. PPARs: therapeutic targets for metabolic disease. *Trends Pharmacol. Sci.* 26 (5), 244–251.
- Bijland, S., et al., 2011. Perfluoroalkyl sulfonates cause alkyl chain length-dependent hepatic steatosis and hypolipidemia mainly by impairing lipoprotein production in APOE\*3-Leiden CETP mice. *Toxicol. Sci.* 123 (1), 290–303.
- Bocher, V., et al., 2002. PPARs: transcription factors controlling lipid and lipoprotein metabolism. *Ann. N. Y. Acad. Sci.* 967, 7–18.
- Boujrad, N., et al., 2000. The peroxisome proliferator perfluorodecanoic acid inhibits the peripheral-type benzodiazepine receptor (PBR) expression and hormone-stimulated mitochondrial cholesterol transport and steroid formation in Leydig cells. *Endocrinology* 141 (9), 3137–3148.
- Brunius, C., Shi, L., Landberg, R., 2016. Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction. *Metabolomics* 12 (11), 173.
- Calafat, A.M., et al., 2007. Polyfluoroalkyl chemicals in the U.S. population: data from the National Health and Nutrition Examination Survey (NHANES) 2003–2004 and comparisons with NHANES 1999–2000. *Environ. Health Perspect.* 115 (11), 1596–1602.
- Cardenas, A., et al., 2018. Association of perfluoroalkyl and polyfluoroalkyl substances with adiposity. *JAMA Netw Open* 1 (4), e181493.
- Chadeau-Hyam, M., et al., 2011. Meeting-in-the-middle using metabolic profiling – a strategy for the identification of intermediate biomarkers in cohort studies. *Biomarkers* 16 (1), 83–88.
- Conder, J.M., et al., 2008. Are PFCA's Bioaccumulative? A critical review and comparison with regulatory criteria and persistent lipophilic compounds. *Environ. Sci. Technol.* 42 (4), 995–1003.
- Conway, B., Innes, K.E., Long, D., 2016. Perfluoroalkyl substances and beta cell deficient diabetes. *J. Diabetes Compl.* 30 (6), 993–998.
- Dambrova, M., Liepinsh, E., 2015. Risks and benefits of carnitine supplementation in diabetes. *Exp. Clin. Endocrinol. Diabetes* 123 (02), 95–100.
- Donat-Vargas, C., et al., 2019a. Perfluoroalkyl substances and risk of type II diabetes: a prospective nested case-control study. *Environ. Int.* 123, 390–398.
- Donat-Vargas, C., et al., 2019b. Associations between repeated measure of plasma perfluoroalkyl substances and cardiometabolic risk factors. *Environ. Int.* 124, 58–65.
- Fievet, C., Fruchart, J.C., Staels, B., 2006. PPARalpha and PPARgamma dual agonists for the treatment of type 2 diabetes and the metabolic syndrome. *Curr. Opin. Pharmacol.* 6 (6), 606–614.
- Gängler, S., et al., 2019. Exposure to disinfection byproducts and risk of type 2 diabetes: a nested case-control study in the HUNT and Lifelines cohorts. *Metabolomics* 15 (4), 60.
- Hanhineva, K., et al., 2014. Nontargeted metabolite profiling discriminates diet-specific biomarkers for consumption of whole grains, fatty fish, and bilberries in a randomized controlled trial. *J. Nutr.* 145 (1), 7–17.
- Haug, L.S., et al., 2010. Diet and particularly seafood are major sources of perfluorinated compounds in humans. *Environ. Int.* 36 (7), 772–778.
- Haug, L.S., et al., 2011. Characterisation of human exposure pathways to perfluorinated compounds — comparing exposure estimates with biomarkers of exposure. *Environ. Int.* 37 (4), 687–693.
- He, X., et al., 2018. PFOA is associated with diabetes and metabolic alteration in US men: National Health and Nutrition Examination Survey 2003–2012. *Sci. Total Environ.* 625, 566–574.
- Hu, X., et al., 2019. Metabolome wide association study of serum poly and perfluoroalkyl substances (PFASs) in pregnancy and early postpartum. *Reprod. Toxicol.* 87, 70–78.
- Jannasch, F., Kroger, J., Schulze, M.B., 2017. Dietary patterns and type 2 diabetes: a systematic literature review and meta-analysis of prospective studies. *J. Nutr.* 147 (6), 1174–1182.
- Jensen, A.A., Leffers, H., 2008. Emerging endocrine disruptors: perfluoroalkylated substances. *Int. J. Androl.* 31 (2), 161–169.
- Johansson, G., et al., 2001. Underreporting of energy intake in repeated 24-hour recalls related to gender, age, weight status, day of interview, educational level, reported food intake, smoking habits and area of living. *Public Health Nutr.* 4 (4), 919–927.
- Johansson, I., et al., 2002. Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. *Public Health Nutr.* 5 (3), 487–496.
- Kar, S., et al., 2017. Endocrine-disrupting activity of per- and polyfluoroalkyl substances: Exploring combined approaches of ligand and structure based modeling. *Chemosphere* 184, 514–523.
- Karnes, C., Winquist, A., Steenland, K., 2014. Incidence of type II diabetes in a cohort with substantial exposure to perfluorooctanoic acid. *Environ. Res.* 128, 78–83.
- Kees, K.L., et al., 1992. Perfluorocarbon-based antidiabetic agents. *J. Med. Chem.* 35 (5), 944–953.
- Kim, S.J., et al., 2016. Understanding metabolomics in biomedical research. *Endocrinol. Metab. (Seoul)* 31 (1), 7–16.
- Kingsley, S.L., et al., 2019. Metabolomics of childhood exposure to perfluoroalkyl substances: a cross-sectional study. *Metabolomics* 15 (7), 95.
- Koeth, R.A., et al., 2014.  $\gamma$ -Butyrobetaine is a proatherogenic intermediate in gut microbial metabolism of L-carnitine to TMAO. *Cell Metab.* 20 (5), 799–812.
- Koeth, R.A., et al., 2019. L-Carnitine in omnivorous diets induces an atherogenic gut microbial pathway in humans. *J. Clin. Investig.* 129 (1), 373–387.
- Koponen, J., et al., 2013. Determination of selected perfluorinated alkyl acids and persistent organic pollutants from a small volume human serum sample relevant for epidemiological studies. *J. Chromatogr. A* 1309, 48–55.
- Lau, C., et al., 2007. Perfluoroalkyl acids: a review of monitoring and toxicological findings. *Toxicol. Sci.* 99 (2), 366–394.
- Lee, J.E., Choi, K., 2017. Perfluoroalkyl substances exposure and thyroid hormones in humans: epidemiological observations and implications. *Ann. Pediatr. Endocrinol. Metab.* 22 (1), 6–14.
- Lefebvre, P., et al., 2006. Sorting out the roles of PPAR alpha in energy metabolism and vascular homeostasis. *J. Clin. Invest.* 116 (3), 571–580.
- Li, C.-H., et al., 2019. Investigation of binding and activity of perfluoroalkyl substances to the human peroxisome proliferator-activated receptor  $\beta/\delta$ . *Environ. Sci.: Process. Impacts*.
- Li, D., Kirsop, J., Tang, W.H.W., 2015. Listening to our gut: contribution of gut microbiota and cardiovascular risk in diabetes pathogenesis. *Curr. Diab. Rep.* 15 (9).
- Lindgren, F., et al., 1996. Model validation by permutation tests: applications to variable selection. *J. Chemom.* 10 (5–6), 521–532.
- Liu, G., et al., 2018. Perfluoroalkyl substances and changes in body weight and resting metabolic rate in response to weight-loss diets: a prospective study. *PLoS Med.* 15 (2), e1002502.
- MacNeil, J., et al., 2009. A cross-sectional analysis of type II diabetes in a community with exposure to perfluorooctanoic acid (PFOA). *Environ. Res.* 109 (8), 997–1003.
- Maillard, E., et al., 2008. Perfluorocarbons: new tool for islets preservation in vitro. *Transplant Proc.* 40 (2), 372–374.
- Maillard, E., et al., 2012. Perfluorocarbon emulsions prevent hypoxia of pancreatic beta-cells. *Cell Transplant.* 21 (4), 657–669.
- Namazi, N., et al., 2019. The association between types of seafood intake and the risk of type 2 diabetes: a systematic review and meta-analysis of prospective cohort studies. *Health Promot. Perspect.* 9 (3), 164–173.
- Nettleton, J.A., et al., 2013. Meta-analysis investigating associations between healthy diet and fasting glucose and insulin levels and modification by loci associated with glucose homeostasis in data from 15 cohorts. *Am. J. Epidemiol.* 177 (2), 103–115.
- Norberg, M., et al., 2010. The Vasterbotten Intervention Programme: background, design and implications. *Glob. Health Action* 3.
- Olokoba, A.B., Obateru, O.A., Olokoba, L.B., 2012. Type 2 diabetes mellitus: a review of current trends. *Oman Med J* 27 (4), 269–273.
- Olsen, G.W., et al., 2007. Half-life of serum elimination of perfluorooctanesulfonate, perfluorohexanesulfonate, and perfluorooctanoate in retired fluorocarbon production workers. *Environ. Health Perspect.* 115 (9), 1298–1305.
- Poynter, M.E., Daynes, R.A., 1998. Peroxisome proliferator-activated receptor alpha activation modulates cellular redox status, represses nuclear factor-kappaB signaling, and reduces inflammatory cytokine production in aging. *J. Biol. Chem.* 273 (49), 32833–32841.

- R Core Team, 2017. R: A language and environment for statistical computing. In: R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Ratnayake, N.J.W., et al., 2018. Beyond genomics: understanding exposotypes through metabolomics. *Human Genomics* 12 (1), 4.
- Razquin, C., et al., 2018. Plasma lipidomic profiling and risk of type 2 diabetes in the PREDIMED trial. *Diabetes Care* 41 (12), 2617–2624.
- Richardson, D.B., Ciampi, A., 2003. Effects of exposure measurement error when an exposure variable is constrained by a lower limit. *Am. J. Epidemiol.* 157 (4), 355–363.
- Ringseis, R., Wen, G., Eder, K., 2012. Regulation of genes involved in carnitine homeostasis by PPAR $\alpha$  across different species (rat, mouse, pig, cattle, chicken, and human). *PPAR Res.* 2012, 868317.
- Rolandsson, O., et al., 2012. How to diagnose and classify diabetes in primary health care: lessons learned from the Diabetes Register in Northern Sweden (DiabNorth). *Scand. J. Prim. Health Care* 30 (2), 81–87.
- Rosen, M.B., et al., 2017. PPAR $\alpha$ -independent transcriptional targets of perfluoroalkyl acids revealed by transcript profiling. *Toxicology* 387, 95–107.
- Salihovic, S., et al., 2019. Identification of metabolic profiles associated with human exposure to perfluoroalkyl substances. *J. Expo Sci. Environ. Epidemiol.* 29 (2), 196–205.
- Saltiel, A.R., Kahn, C.R., 2001. Insulin signalling and the regulation of glucose and lipid metabolism. *Nature* 414 (6865), 799–806.
- Schillemans, T., et al., 2019. Visualization and interpretation of multivariate associations with disease risk markers and disease risk-the triplot. *Metabolites* 9 (7).
- Sheng, N., et al., 2018. Cytotoxicity of novel fluorinated alternatives to long-chain perfluoroalkyl substances to human liver cell line and their binding capacity to human liver fatty acid binding protein. *Arch. Toxicol.* 92 (1), 359–369.
- Shi, Z., et al., 2007. Alterations in gene expression and testosterone synthesis in the testes of male rats exposed to perfluorododecanoic acid. *Toxicol. Sci.* 98 (1), 206–215.
- Shi, Z., et al., 2010. Perfluorododecanoic acid-induced steroidogenic inhibition is associated with steroidogenic acute regulatory protein and reactive oxygen species in cAMP-stimulated Leydig cells. *Toxicol. Sci.* 114 (2), 285–294.
- Shi, L., et al., 2017. Targeted metabolomics reveals differences in the extended postprandial plasma metabolome of healthy subjects after intake of whole-grain rye porridges versus refined wheat bread. *Mol. Nutr. Food Res.* 61 (7).
- Shi, L., et al., 2018a. Plasma metabolites associated with type 2 diabetes in a Swedish population: a case-control study nested in a prospective cohort. *Diabetologia* 61 (4), 849–861.
- Shi, L., et al., 2018b. Plasma metabolites associated with healthy Nordic dietary indexes and risk of type 2 diabetes-a nested case-control study in a Swedish population. *Am. J. Clin. Nutr.*
- Shi, L., et al., 2019a. Variable selection and validation in multivariate modelling. *Bioinformatics* 35 (6), 972–980.
- Shi, L., et al., 2019b. Joint analysis of metabolite markers of fish intake and persistent organic pollutants in relation to type 2 diabetes risk in Swedish adults. *J. Nutr.* 149 (8), 1413–1423.
- Strand, E., et al., 2018. Serum carnitine metabolites and incident type 2 diabetes mellitus in patients with suspected stable angina pectoris. *J. Clin. Endocrinol. Metab.* 103 (3), 1033–1041.
- Strynar, M.J., Lindstrom, A.B., 2008. Perfluorinated compounds in House Dust from Ohio and North Carolina, USA. *Environ. Sci. Technol.* 42 (10), 3751–3756.
- Stubleski, J., et al., 2017. The effect of drinking water contaminated with perfluoroalkyl substances on a 10-year longitudinal trend of plasma levels in an elderly Uppsala cohort. *Environ. Res.* 159, 95–102.
- Sun, Q., et al., 2018. Plasma concentrations of perfluoroalkyl substances and risk of type 2 diabetes: a prospective investigation among U.S. Women. *Environ. Health Perspect.* 126 (3), 037001.
- Sunderland, E.M., et al., 2019. A review of the pathways of human exposure to poly- and perfluoroalkyl substances (PFASs) and present understanding of health effects. *J. Exposure Sci. Environ. Epidemiol.* 29 (2), 131–147.
- Tornevi, A., et al., 2019. Chlorinated persistent organic pollutants and type 2 diabetes - a population-based study with pre- and post- diagnostic plasma samples. *Environ. Res.* 174, 35–45.
- Vanden Heuvel, J.P., et al., 2006. Differential activation of nuclear receptors by perfluorinated fatty acid analogs and natural fatty acids: a comparison of human, mouse, and rat peroxisome proliferator-activated receptor- $\alpha$ , - $\beta$ , and - $\gamma$ , liver X receptor- $\beta$ , and retinoid X receptor- $\alpha$ . *Toxicol. Sci.* 92 (2), 476–489.
- Vestergren, R., Cousins, I.T., 2009. Tracking the pathways of human exposure to perfluorocarboxylates. *Environ. Sci. Technol.* 43 (15), 5565–5575.
- Vineis, P., Perera, F., 2007. Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old. *Cancer Epidemiol. Biomarkers Prev.* 16 (10), 1954–1965.
- Walker, D.L., et al., 2019. The metabolome: a key measure for exposome research in epidemiology. *Curr. Epidemiol. Rep.* 6 (2), 93–103.
- Wild, C.P., 2005. Complementing the genome with an “exposome”: the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomark. Prev.* 14 (8), 1847–1850.
- Wolf, C.J., et al., 2008a. Activation of mouse and human peroxisome proliferator-activated receptor  $\alpha$  by perfluoroalkyl acids of different functional groups and chain lengths. *Toxicol. Sci.* 106 (1), 162–171.
- Wolf, D.C., et al., 2008b. Comparative hepatic effects of perfluorooctanoic acid and WY 14,643 in PPAR- $\alpha$  knockout and wild-type mice. *Toxicol. Pathol.* 36 (4), 632–639.
- Yan, S., et al., 2015. Perfluorooctanoic acid exposure for 28 days affects glucose homeostasis and induces insulin hypersensitivity in mice. *Sci. Rep.* 5, 11029.
- Zhao, B., et al., 2010. Inhibition of human and rat 3 $\beta$ -hydroxysteroid dehydrogenase and 17 $\beta$ -hydroxysteroid dehydrogenase 3 activities by perfluoroalkylated substances. *Chem. Biol. Interact.* 188 (1), 38–43.